

## Predicting new venture survival and growth: does the fog lift?

Article (Published Version)

Coad, Alex, Frankish, Julian S, Roberts, Richard G and Storey, David J (2016) Predicting new venture survival and growth: does the fog lift? *Small Business Economics*, 47 (1). pp. 217-241. ISSN 0921-898X

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/60555/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

# Predicting new venture survival and growth: Does the fog lift?

Alex Coad · Julian S. Frankish ·  
Richard G. Roberts · David J. Storey

Accepted: 10 February 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** This paper investigates whether new venture performance becomes easier to predict as the venture ages: does the fog lift? To address this question we primarily draw upon a theoretical framework, initially formulated in a managerial context by Levinthal (Adm Sci Q 36(3):397–420, 1991) that sees new venture sales as a random walk but survival being determined by the stock of available resources (proxied by size). We derive theoretical predictions that are tested with a 10-year cohort of 6579 UK new ventures in the UK. We observe that our ability to predict firm growth deteriorates in the years after entry—in terms of the selection environment, the ‘fog’ seems to thicken. However, our survival predictions improve with time—implying that the ‘fog’ does lift.

**Keywords** Entrepreneurship · Firm growth · Survival analysis · Coefficient of determination · Selection environment · Gambler’s Ruin theory

**JEL classifications** L26 · L25

## 1 Introduction

Economic dynamics are characterised by a noisy selection environment that (imperfectly) rewards superior performance. This paper investigates whether the selection environment becomes clearer—more predictable—in the years after entry. The benefits of greater predictability accrue to business owners, to providers of finance and to governments. For business owners, the value is the availability of a route-map to enable them to plan ahead and check progress over time (Dencker et al. 2009). For providers of finance, being able to more accurately estimate the optimal date to provide finance is valuable, because too early an investment may be too risky, whereas delay may mean the opportunity is seized by a rival (Cumming et al. 2015). Finally, governments are continually faced with the choice of using taxpayer’s funds to support and stimulate start-ups, or instead to delay support until performance metrics become clearer (Pons Rotger et al. 2012). An optimal combination of support at different stages as new ventures evolve could provide considerable social and economic returns.

---

A. Coad (✉) · D. J. Storey  
University of Sussex, Jubilee Building, Falmer,  
Brighton BN19SL, UK  
e-mail: A.Coad@sussex.ac.uk

D. J. Storey  
e-mail: D.J.Storey@sussex.ac.uk

A. Coad  
JRC-IPTS, European Commission, Edificio Expo,  
41092 Seville, Spain

J. S. Frankish  
Barclays Bank, Poole, Dorset, UK

R. G. Roberts  
University of Birmingham, Birmingham, UK

This paper is motivated by a desire to explain and then to use these explanations to predict, post-start performance so providing the benefits of greater predictability to all three parties. It takes two alternative measures of performance (Miller et al. 2013)—survival and sales growth—and assesses whether, over time, our ability to explain these two performance variables improves. In the phrasing of this paper—does the fog lift with time? If so, when does this greater clarity appear? Is it after 1 year, or 10 years? Or, does the fog lift only gradually but continually? Is there a clear ‘step’ at a clearly-identified point in time?

Our theoretical starting point is the Levinthal (1991) random walk model which we apply to new, as opposed to well-established, ventures. Here, each enterprise has an initial stock of resources which expands or contracts depending on post-entry growth, which is determined by stochastic shocks. Exit takes place when the stock falls below a minimum threshold.<sup>1</sup>

Given these assumptions we distinguish between new venture sales growth and survival. Following Levinthal (1991), if the sales growth of a new venture follows a random walk, there is no improvement in our ability to predict growth in the years after entry—hence the fog is thick and remains thick over time. However, we assume new ventures have different (financial) resource endowments, enabling those with more resources to survive shocks that would lead to the exit of those with fewer resources. These financial endowments are either present at start-up, or accumulated through post-entry growth. Survival rates are therefore expected to increase and become more predictable, in the years after entry, as surviving new ventures acquire the financial resources that enable them to ‘ride out’ the inevitable vicissitudes of trade that characterise their early months and years.

These predictions for growth and survival are tested using a cohort of 6579 new ventures in the UK, all of which began to trade in the same quarter of 2004, where *every financial transaction* is tracked over 10 years. With this unique data we show that our

ability to explain sales growth decreases as the venture ages, because as time goes by this becomes more random. When we focus only on firms that survive until the end of year 10; however, for this subsample of surviving firms, our ability to predict growth remains constant over time. Regarding survival, our ability to predict which firms will remain in operation increases slightly in the years after entry. Our results are therefore broadly consistent with our model.

Our specific contribution is then to demonstrate that, even if the sales growth of a new venture increasingly approximates a random walk, its survival becomes more predictable. The growth fog becomes thicker over time, but the survival fog becomes less dense. Perhaps the paper most closely related to ours is Lotti et al. (2009), who present evidence that firms converge to a random growth model (i.e. Gibrat’s (1931) ‘Law of Proportionate Effect’) in the years after entry. In our paper, however, we look more widely at our ability to explain growth and survival in the years after entry. Another related paper is Wiklund et al. (2010), who observe that the explanatory power of financial indicators decreases in the years after entry, when the task is to explain survival. In our analysis, we include other variables (beyond financial indicators) as explanatory variables for performance (measured in terms of both survival and growth), and present finer-grained evidence on the year-by-year evolution of the model fit statistics.

The remainder of the paper is set out as follows: Sect. 2 provides the theoretical context that is used in Sect. 3 to derive hypotheses. Section 4 presents our methodology. Section 5 presents the dataset, and we test our hypotheses in Sect. 6. Section 7 concludes.

## 2 Theory development

Conceptualising firm performance as a random walk has a long history in economics (Gibrat 1931; Ijiri and Simon 1964; Levinthal 1991; Denrell et al. 2015). Random processes produce results that closely match the outcomes of many top performing companies, to the extent that investigating whether or not performance is purely random remains a valid research question (Henderson et al. 2012; Denrell et al. 2015; Storey 2011). This need not imply that managers do not put thoughtful planning and effort into their business decisions, because it could be that competition is so

<sup>1</sup> Gimeno et al. (1997) extend this framework to allow for individual-specific thresholds, according to which some individuals with attractive outside options may exit before they reach a minimum level of resources. Furthermore, one could conceive possible extensions in which the exit threshold is not exogenous but endogenous and potentially time-variant.

fierce, and businesses are all more or less ‘neck-and-neck’, that there may not be any easily observed systematic factors that allow new ventures to enjoy prolonged above-average performance in the years after entry. ‘Chance models are, in fact, compatible with effortful managers who carry out deliberate actions’ (Denrell et al. 2015, p. 936). Our preference for chance models in this paper is because random walk models offer useful approximations to real-world phenomena (Levinthal 1991; Henderson et al. 2012; Denrell et al. 2015), and also because random walk models can provide simple and clear theoretical predictions that can be developed into testable hypotheses.

Levinthal (1991) was amongst the first to formally explore how random processes could shed light on venture survival in a managerial context. His model had two key assumptions. The first was that firm growth was modelled as a random walk, and the second was that survival depended upon access to resources or assets that could be used to finance the shocks experienced by the business in a random walk.<sup>2</sup>

Levinthal emphasised that the random walk model is compatible with variations in competence amongst enterprises. He writes (p. 399):

While variation in competence should shift the mean of the possible distribution of outcomes, and perhaps the variance as well, the presence or absence of competence does not fundamentally alter the stochastic nature of the process.

The Levinthal (1991) model also puts forward that the amount of the assets is determined by two factors—past performance and initial resources. We assume that access to more financial resources improves the chances of survival. However, there are two key respects in which the assets of the new venture differ from that in an established firm. The first is that, in an established venture, the assets primarily comprise those accumulated over time, whereas those available to the new venture are considerably more likely to be those in place when the venture begins. Second, in an established firm the accumulated assets constitute a

‘track-record’ which can help internal and external parties assess future performance, whereas no such record exists for a new venture. This is particularly problematic for external suppliers of finance—banks, trade creditors—who then seek ‘signals’ of credibility, such as collateral (Voordeckers and Steijvers 2006).

Our model therefore assumes the returns from venture creation are a random walk and this payoff structure attracts individuals who are optimistic and favour situations where, although the expected returns may be negative (Hamilton 2000), the variance is high and positively skewed. Survival, in turn, reflects the availability of resources (i.e. resources available at start-up, as well as those obtained from post-entry performance).

More formally expressed, growth occurs through the following random process:

$$x_t = x_{t-1} + \varepsilon_t, \quad (1)$$

where  $x_t$  is the logarithm of firm size at time  $t$ , and  $\varepsilon$  is a random shock (additive in logs, but multiplicative on a linear scale) with mean  $\mu$  and standard deviation  $\sigma$ .

Survival is a function of the stock of accumulated resources, so survival,  $S$ , depends on whether a firm’s resources exceed a minimum threshold size  $x^*$ :

$$S = 1 \quad \text{if } x^l > x^*; \quad \text{otherwise } S = 0 \quad (2)$$

where  $x^l$  is a latent variable that corresponds to  $x$  if  $x^l > x^*$ , but remains unobserved if  $x^l \leq x^*$ . If the exit threshold is positive, i.e.  $x^* > 0$ , then players will not persist until their resources reach zero, but quit the ‘gambling table’ even when resources are positive (Gimeno et al. 1997).

### 3 Hypotheses derivation

Our primary interest is in whether the selection environment for new ventures improves—becomes more predictable—in the years after entry. To do this we investigate the explanatory power (or goodness-of-fit, represented by the  $R^2$  statistic) of models that seek to explain the growth and survival of new ventures.

#### 3.1 Growth

If new venture growth is a random walk, à la Levinthal, the dynamics of (log) size are  $x_t = x_{t-1} + \varepsilon_t$ , the growth rate (in log-differences; Tornqvist et al. 1985) is

<sup>2</sup> The Levinthal (1991) model assumes that exit takes place when the business cannot meet its financial obligations. However, Gimeno et al (1997) shows that the ability to assemble these resources is endogenous in the sense that they vary according to the alternative employment options open to the business owner(s).

expressed entirely in terms of a random shock:  $x_t - x_{t-1} = \varepsilon_t$ . Growth is well approximated by a random walk in the years after start-up, and our inability to make systematic predictions for post-entry growth implies that the expected  $R^2$  from growth regressions is low and remains low in the years after entry.<sup>3</sup>

**Hypothesis 1** The  $R^2$  from regressions of the determinants of new venture growth does not increase in the years after entry.

### 3.2 Survival

To examine survival, firm size at time  $t$  is  $x_t$ , with start-up size being denoted as  $x_0$ . In a random walk model à la Levinthal (1991), firm size evolves, with  $x_t = x_{t-1} + \varepsilon_t$ , where  $\varepsilon_t$  is distributed with mean  $\mu$  and variance  $\sigma^2$ . When  $\mu = 0$ , we have a pure random walk, whereas when  $\mu > 0$  (following Le Mens et al. 2011) then there is a steady increase in expected resource stock over time.

Firms are assumed to exit when their size (proxied by their resource stock) reaches zero. The time taken until the firm first exhausts its resources (i.e.  $x^j \leq x^*$ , for the case when  $x^* = 0$ ) is expressed as the cumulative distribution function of a random variable in the following way (known as the Bachelier–Lévy formula):

$$F(T|x_0, \mu, \sigma) = N\left(-\frac{\mu t + x_0}{\sigma\sqrt{t}}\right) + e^{-2x_0\frac{\mu}{\sigma^2}} \cdot N\left(\frac{\mu t - x_0}{\sigma\sqrt{t}}\right) \quad (3)$$

where  $N(\cdot)$  represents the cumulative density of the standard normal distribution (Levinthal 1991; Coad et al. 2014). Time to exit is thus a function of three parameters: the trend in the random walk  $\mu$ , the variance  $\sigma^2$  of the growth shocks, and start-up size  $x_0$ .<sup>4</sup>

<sup>3</sup> We consider it trivial that the  $R^2$  will be low and driven by stochastic noise, therefore we do not see the need to use a simulation model here to demonstrate the evolution of the  $R^2$ .

<sup>4</sup> It was correctly pointed out to us by a Referee that this does not capture the situation where “a wealthy entrepreneur operates a very small firm (with a low amount of annual sales). Such a small firm could experience negative shocks which are then financed by the personal wealth of the entrepreneur”. We acknowledge that this highlights a mismatch between the theoretical construct and the empirics. However we do not see how this could be resolved, since even the Bank has an imperfect idea of the wealth of its clients and would only be incentivised to quantify that wealth in the unlikely event that this particular NV were seeking large funding.

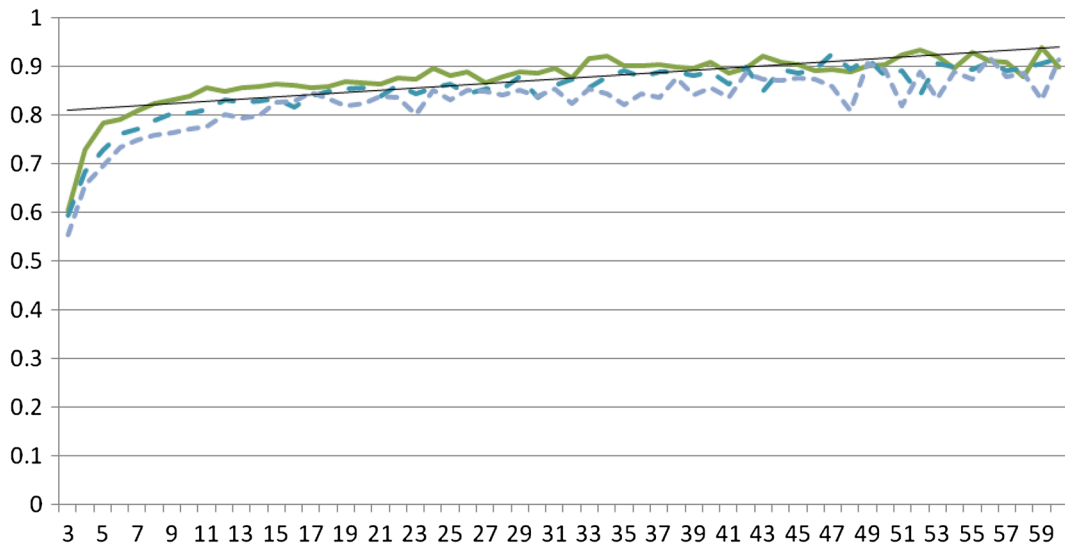
Even if growth is a random process, expected survival time can be increased by increasing the size at start-up  $x_0$  (Levinthal 1991; Coad et al. 2014). The  $R^2$  from survival regressions therefore depends on both start-up size and growth since start-up.

We now apply a simulation model to derive implications of the Levinthal random walk model for the evolution of the  $R^2$ . We generate an artificial dataset of 50,000 firms, whose start-up size is calibrated according to the lognormal distribution with mean 10.55 and standard deviation 1.5, in order to closely follow the start-up size distribution observed in our data. We then generate a distribution of growth rates, distributed according to the Laplace or ‘symmetric exponential’ (Stanley et al. 1996; Bottazzi and Secchi 2006), with mean  $\mu = -0.1$  and standard deviation  $\sigma = 0.9$  (again, closely following the values observed in our data).<sup>5</sup> Firm size evolves as a random walk,  $x_t = x_{t-1} + \varepsilon_t$ , given the distributions of start-up size and growth rates given above, for  $t = 60$  periods. The exit threshold  $x^*$  is set at 7 in the baseline case, which is deliberately chosen to be a relatively high value that will guarantee that in each period some firms will exit (thus avoiding a degenerate value for the  $R^2$  in any year’s survival regression in which all firms survive). For each individual period up to  $t = 60$ , we estimate a probit survival regression (with a constant term and a single explanatory variable: lagged size) and record the Nagelkerke  $R^2$  statistic.

Figure 1 shows that the  $R^2$  clearly increases in the years following start-up. This is because, with the passage of time, surviving firms overcome the liability of newness and grow to become sufficiently large that they have accumulated a ‘buffer’ stock of resources, and no longer operate on the brink of the exit threshold. Firms that start small, on the other hand, are more likely to be quickly weeded out through a selection effect. As these chaotic, short-lived firms are removed, the selection environment becomes less ‘foggy’. The central point here is that the  $R^2$  value rises over time even when growth is a random walk.

**Hypothesis 2** The  $R^2$  from regressions of the determinants of new venture survival increases in the years after entry.

<sup>5</sup> See Table 1 for summary statistics on our dataset.



**Fig. 1** Evolution of the Nagelkerke  $R^2$  using simulated data, for 60 periods. y-axis: Nagelkerke  $R^2$  obtained from probit regressions where exit depends on lagged size. x-axis: time

period. *Baseline* case (with exit threshold  $x^* = 7$ ) appears as a *solid line*;  $x^* = 8$  for the *long-dash line*;  $x^* = 9$  for the *short-dash line*. Linear trend-line plotted for the *baseline case*

#### 4 Testing for changes in the density of the fog

Of crucial interest for our paper is measuring what we call ‘fog’—the coefficient of determination, or  $R^2$  statistic. The standard  $R^2$  statistic is expressed in terms of how well an OLS regression model can explain the total variation in the data:

$$R^2 \equiv \frac{SS_{\text{reg}}}{SS_{\text{tot}}} = 1 - \left\{ \frac{SS_{\text{res}}}{SS_{\text{tot}}} \right\}$$

where  $SS_{\text{reg}}$  is the regression sum of squares (i.e. the explained sum of squares),  $SS_{\text{tot}}$  is the total sum of squares, and  $SS_{\text{res}}$  is the residual (i.e. unexplained) sum of squares. The  $R^2$  statistic provides meaningful information on how well a set of variables can explain a given outcome, or how well we can predict real-world outcomes on the basis of our available information (Bertrand and Schoar 2003; see also Syverson 2011, p. 340). Cox and Snell (1989) suggested that the  $R^2$  statistic be generalised to other regression models (such as regression models with binary dependent variables) where maximum likelihood is the criterion of fit. They suggested the following  $R^2$  statistic:

$$\text{Cox-Snell } R^2 = 1 - \left\{ L(0)/L(\hat{\beta}) \right\}^{\frac{2}{n}}$$

where  $L(\hat{\beta})$  and  $L(0)$  denote the likelihoods of the fitted and ‘null’ models, respectively. The Cox–Snell  $R^2$

statistic has a number of desirable properties (e.g. it is asymptotically independent of the sample size), although a drawback is that it reaches a maximum value that is lower than unity for discrete models (Nagelkerke 1991). Therefore, it has been suggested that the Cox–Snell  $R^2$  be adjusted as follows, to obtain what has become known as the Nagelkerke  $R^2$  statistic, after Nagelkerke (1991):

$$\text{Nagelkerke } R^2 = \text{Cox-Snell } R^2 / [\max(R^2)]$$

$$\text{where } \max(R^2) = 1 - L(0)^{\frac{2}{n}}.$$

Because of its desirable statistical properties we use the Nagelkerke  $R^2$  statistic, although we check that our results are not sensitive to this choice of  $R^2$  statistic.

We begin by running regressions on cross sections corresponding to each year, where the dependent variable is either growth rate or survival probability.<sup>6</sup> For each year we obtain a Nagelkerke  $R^2$  statistic. We then plot the evolution of the Nagelkerke  $R^2$  over time using line charts—one chart for growth, one for survival.

<sup>6</sup> Note that, in contrast to a large body of research on survival models, we do not investigate the determinants of a firm’s total survival *duration*, but instead the chances of survival for any one particular year.



## 5 The dataset: Barclays bank customer accounts

### 5.1 Start-up: definition

We exploit a rich and unique dataset drawn from non-financial firms identified as start-ups or new ventures that entered the business customer base of Barclays Bank between March and May 2004. At that time about one in four UK start-ups banked with Barclays. The sample excludes established businesses that switched from another Bank. We are aware that a new business does not necessarily start trading immediately upon opening an account. Indeed, for Barclays' customers, approximately five per cent of start-ups show no activity through their account in the subsequent 12 months. We addressed this by only including firms that showed activity in the month following entry to the customer base.<sup>7,8</sup>

We therefore focus on a cohort of 6579 firms that have the same start date. We consider this to be important, because firms starting in different years may not be readily comparable (especially if the macroeconomic conditions at start-up have persistent effects on firm development in subsequent years). Focusing on a single cohort means that firms face the same macro-economic conditions at each year of their development and can therefore be meaningfully compared (Ryder 1965; Anyadike-Danes et al. 2015). We then track the cohort for a maximum of 10 years, a period of time that we consider to be sufficiently long for our purposes, given that over 80 % of the ventures will have exited in that time (Anyadike-Danes and Hart 2014).

### 5.2 Start-up: data

Prior to opening the new business account, data were collected on the founder(s) gender, age, highest level of educational qualifications; prior business experience; previous ownership; and/or ownership amongst immediate family members. Finally, to capture access

to non-financial resources, owners were asked about the sources of advice and support they used prior to start-up.

These data were then supplemented by the bank as part of its general account opening process. This covers the legal form of the business, the activity type (sector/branch/market) and its location (standard region) within the UK. Table 2 in 'Appendix' sets out the data definitions in full.

### 5.3 Ongoing data

To measure the size of the business we used credit turnover—the value of payments into a current account,<sup>9</sup> which we will refer to as 'sales'. This serves as a very close approximation to sales revenue inclusive of taxes.<sup>10</sup> The much greater granularity of sales compared with using measures of employee numbers is a particular strength.<sup>11</sup> It is also reliable, comprehensive and, *because every financial transaction is documented*, the scale of volatility can be reliably quantified. Although credit turnover was initially observed by the bank at monthly intervals, the data we have have been aggregated over 12 months to analyse annual values, since our focus here is to explain long-run, rather than short-run changes.

### 5.4 Exit and closure

Establishing precisely when a business has closed is perhaps the most challenging aspect of any study of new ventures. Even for datasets taken from near comprehensive official sources, the date at which exit occurs may be some time after actual closure.<sup>12</sup>

<sup>7</sup> We also included a small proportion of firms who did not show activity in their first full month, but in either May or June 2004. In these cases the start month of the firm was recorded as the month prior to activity.

<sup>8</sup> The UK, unlike many countries in continental Europe, is not characterised by multiple banking (Ongena and Smith 2000). The account at a single bank is therefore likely to capture the full trading activities of the new venture.

<sup>9</sup> Excluding payments from related accounts, e.g. deposit accounts held by the business.

<sup>10</sup> Prior empirical work has measured growth in terms of numerous metrics such as employment, sales, profits, business valuation. We follow Zimmerman and Zeltz (2002, p. 417) who explain that "growth in sales is especially important for new ventures since their economies of scale are too important for them to continue without increasing their scale of operations."

<sup>11</sup> These strengths are discussed in more detail shortly.

<sup>12</sup> For example Storey et al (1987, p. 45), in a study of the closure of 177 Limited Companies that "failed", identified seven decision-rules that were required to identify the year in which the enterprise ceased to trade.

When using bank records, there are two main issues to resolve. The first is to distinguish between those businesses that have closed, and those that have switched to another bank. For our dataset we used Barclays closure-reason-codes that record why any given account has been closed. 1.38 % of our initial sample switched over the 10 years covered by the dataset, i.e. they had closed their account with Barclays, but continued to trade.<sup>13</sup> These were dropped from our sample before we started the analysis.

The second issue is judging when a given business has actually closed. While the majority of Barclays customers ceasing to trade clearly close at a specific time when no more transactions take place, an important minority become dormant, i.e. their account remains open, but with no activity.<sup>14</sup> For the firms in our sample we used a simple rule—if the business had shown no sales in consecutive 6-month periods, then it was deemed to have closed in the first of these periods.<sup>15</sup>

It is important to note that this process identifies closures. It is not limited to business ‘failures’. By the latter we mean those firms that cease to trade with some external financial liability. Of course, as noted earlier, a closing firm may, or may not, have met the objectives of its owner(s), although closure may equally reflect that a better opportunity has presented itself to the business owner(s) (Headd 2003; Harada 2007). Finally, cases of entrepreneurial exit (but business continuation) such as an initial public offering (IPO), merger or acquisition (M&A) or trade sale will not have a confounding effect on our measurement of business exit (Wennberg et al. 2010; Coad 2014), because if the firm continues operations with the same bank account, it will be treated in our dataset as a continuing firm, whereas if it switches its bank account to a different bank, it will be treated in our dataset as a ‘switcher’ and dropped prior to analysis.

<sup>13</sup> This could be an understatement of the true number if there were imperfections in the reporting process meaning that some switchers were not recorded. However, our rates are broadly in line with Fraser (2005, p. 90) for all types of UK (SME) businesses.

<sup>14</sup> Indeed, some of these may have switched rather than closed.

<sup>15</sup> Some Barclays customers can show little or no activity for a number of months before seeing turnover return to non-negligible levels. This reflects the nature of many ‘micro’ businesses.

Nevertheless, cases of IPOs, M&As and even trade sales are negligible because our new ventures are both young, small and representative of all sectors (apart from financial services). The tech-based services in which these outcomes are particularly characteristic constitute only a tiny proportion of the sample.<sup>16</sup>

#### 5.4.1 Dependent variables

We take two dependent variables as alternative indicators of new venture ‘performance’ (Miller et al. 2013). *Survival* is a binary variable, equal to 1 if the enterprise continues to trade at end of period (=0 if the enterprise exited). The *Growth Rate* is measured in terms of growth in credit turnover (or ‘sales’, the value of payments into a current account) excluding payments from a related account (deposit account).<sup>17</sup> Sales growth has many advantages over other metrics of growth, such as employment, for new ventures. The first is because growth in terms of employment is ‘clunky’ (Coad et al. 2015, p. 6) due to integer constraints in terms of employee headcounts. These are particularly important for new ventures (e.g. a solo self-employed individual contemplating her first hire, who can either remain static or double her size—and nothing in between).<sup>18</sup> Second, the decision to take on a new/first employee is a huge decision by a NV and presents problems of interpretation since it reflects a, difficult to specify, combination of past and current performance as well as future expectations. Finally, most new ventures, in our sample, are too small to employ others—certainly when they start to trade.<sup>19</sup>

<sup>16</sup> It has been pointed out to us, both by referees and others, that this procedure of retaining the same bank account number is not the case in other countries and contexts. We think our classification is a clear benefit, since it also includes the tiny proportion of new ventures that might grow by acquisition in their early years.

<sup>17</sup> Our indicator of growth rates is continuous and can be positive or negative (i.e. decline). We agree with Davila et al. (2015) that the research community has focussed disproportionately upon both growth and high growth.

<sup>18</sup> Detailed data on fractional or part-time employees, total wage bill, or total hours worked, are extremely unusual in longitudinal databases on new venture performance.

<sup>19</sup> Datasets which identify new ventures as those taking their first employee are therefore likely to be both larger and longer-established than those identified as making sales for the first time.



We calculate the annual growth rate in the usual way (see e.g. Tornqvist et al. 1985; Coad 2009) by taking log-differences of sales, i.e.

$$\text{Sales growth}(i, t) = (\log(\text{sales}(i, t)) - \log(\text{sales}(i, t - 1))) \quad (4)$$

We estimate regression equations 1 year at a time, one cross section at a time, to obtain an  $R^2$  statistic for each year. A logistic regression model is applied for our survival estimations (Jenkins 1995; Wiklund et al. 2010), which is compatible with our focus on survival/death within a single year (rather than survival durations over many years). Our regression equations for the growth and survival of firm  $i$  in year  $t$  are as follows:

$$\begin{aligned} \text{Growth}(i, t) = & \alpha_1 + \beta_1 \cdot \log\_sales(i, t - 1) \\ & + \beta_2 \cdot \text{Growth}(i, t - 1) \\ & + \gamma_1 \cdot \text{Entrepreneur}(i, t) \\ & + \delta_1 \cdot \text{Business}(i, t) \\ & + \theta_1 \cdot \text{Account}(i, t) + \varepsilon_1(i, t) \end{aligned} \quad (5)$$

$$\begin{aligned} \text{Survival}(i, t) = & \alpha_2 + \beta_3 \cdot \log\_sales(i, t - 1) \\ & + \beta_4 \cdot \text{Growth}(i, t - 1) \\ & + \gamma_2 \cdot \text{Entrepreneur}(i, t) \\ & + \delta_2 \cdot \text{Business}(i, t) \\ & + \theta_2 \cdot \text{Account}(i, t) + \varepsilon_2(i, t) \end{aligned} \quad (6)$$

where our explanatory variables can be grouped together at the entrepreneur level (age, education, business experience, sources of advice), the business level (number and gender of owner(s), legal form, industry, region), and the bank account level (volatility, overdraft behaviour).<sup>20</sup>

#### 5.4.2 Independent variables

The independent variables used in the analysis are defined in Table 2 in ‘Appendix’. It also sets out where these variables have been used in previous work on survival/growth of new/small enterprises and the results obtained. The first group are the ‘usual

suspects’ such as Legal form (Company, Partnership, Sole Trader); Number of owners; Gender; Age (and Age squared); Education level categories; Sources of advice (EABL scheme, Accountant, Solicitor, College, SR Seminar, PYBT scheme, Family, or Other), and a full set of dummies for industry and geographical region.

A second group is information on bank account activity: sales volatility, availability and extent of use of authorised overdraft facilities, and the use and extent of use of unauthorised overdrafts. These variables have not been explored in previous work that seeks to explain firm growth and survival, and so their inclusion can be considered to be a strength of this paper.<sup>21</sup>

Table 2 in ‘Appendix’ does not point to the omission of key variables that might cause our regression equations to be grossly misspecified. Prior work on firm growth generally has low values for the  $R^2$  statistic (usually lower than 15 %, see the survey in Coad 2009, Table 7.1) so although there remains a risk of specification error and omitted variable bias, there are no clear guidelines in the literature as to which (if any) variables or regression specifications would be more appropriate.

#### 5.5 Summary statistics

Table 1 provides an overview of the size and growth of new ventures in our sample which, it will be recalled, all began trading in the second quarter of 2004. The median sales in year 1 (i.e. 2005) are £38,712 which is far smaller than the threshold for value-added tax (VAT) registration (set at £58,000 for the 12 months from 1 April 2004, and rising to £73,000 by 1 April 2011), above which firms start to appear in UK administrative datasets. Around 50 % of new ventures will exit within 3 years of starting to trade, which is similar to that observed from UK

<sup>20</sup> We do not claim to have resolved the endogeneity between the explanatory variables and the performance outcome variables. We also do not claim to have discovered causal effects regarding growth or survival. Instead, we are interested in describing how the  $R^2$  statistics change in the years after entry.

<sup>21</sup> However, their inclusion may introduce endogeneity into the analysis of growth and survival (e.g. although bank account activity may affect survival and growth, future survival and growth prospects may precede or co-evolve with bank account activity). We therefore remind the reader that the coefficient estimates for our variables reflect partial correlations (i.e. associations) rather than causal effects. We also investigate how our results change according to whether or not these bank account variables are included.

**Table 1** Summary statistics for size and growth rates

	Mean	SD	10 %	25 %	Median	75 %	90 %	Obs.
Sales								
Year 1	114,095	508,678	5475	14,687	38,712	103,658	260,652	5524
Year 2	144,319	546,146	5547	16,529	44,524	124,414	323,178	4162
Year 3	168,352	645,409	5222	17,253	47,855	138,347	373,255	3211
Year 4	183,939	542,018	5438	18,532	51,964	158,026	428,499	2593
Year 5	190,217	552,839	5945	17,996	51,168	152,264	451,445	2152
Year 6	192,157	706,588	5239	17,517	47,924	147,866	453,727	1823
Year 7	213,050	938,730	5700	18,475	53,019	161,941	512,618	1604
Year 8	253,250	1,333,538	6668	19,516	58,134	177,112	577,178	1424
Year 9	277,643	1,640,798	6533	19,274	57,258	180,390	595,597	1311
Year 10	300,699	2,046,271	6860	22,673	64,989	196,821	592,880	1208
Sales growth								
Year 1	–	–	–	–	–	–	–	–
Year 2	–0.055	0.940	–0.964	–0.270	0.053	0.356	0.753	4162
Year 3	–0.133	0.946	–1.001	–0.303	0.022	0.240	0.566	3211
Year 4	–0.110	0.864	–0.873	–0.280	0.013	0.226	0.503	2593
Year 5	–0.189	0.907	–0.991	–0.378	–0.067	0.135	0.427	2152
Year 6	–0.221	0.833	–0.864	–0.368	–0.080	0.086	0.359	1823
Year 7	–0.089	0.772	–0.696	–0.207	0.005	0.185	0.475	1604
Year 8	–0.055	0.698	–0.593	–0.198	0.000	0.184	0.458	1424
Year 9	–0.078	0.731	–0.592	–0.222	–0.022	0.147	0.436	1311
Year 10	–0.037	0.678	–0.518	–0.175	0.020	0.203	0.484	1208

Note that there are 6579 firms at the start of year 1

administrative data on new ventures (Anyadike-Danes and Hart 2014).

To investigate the impact of our rich coverage of micro firms, we complement our baseline results with those obtained from restricting our sample of new ventures to those of above-median start-up size. This makes our sample more similar to other work on new ventures that has a disproportionate coverage of larger new ventures (Yang and Aldrich 2012). The significance of this is that, only by year 10, would the median surviving firm from this dataset have had sales sufficient for them to be included in official data.

A second key fact to emerge from the lower section of Table 1 is that (positive) *growth* in sales is by no means the ‘norm’ for new ventures. *The mean growth rate is negative in every single year*, although the median growth rate is only negative in 3 years. The term ‘sales growth,’ when applied to NVs, is for this reason potentially misleading if it is not understood that growth rates can be negative (Davila et al. 2015). Indeed, negative growth rates (i.e. decline) are very common.

Figure 2 presents the growth rate distribution, which resembles the usual Laplace or symmetric

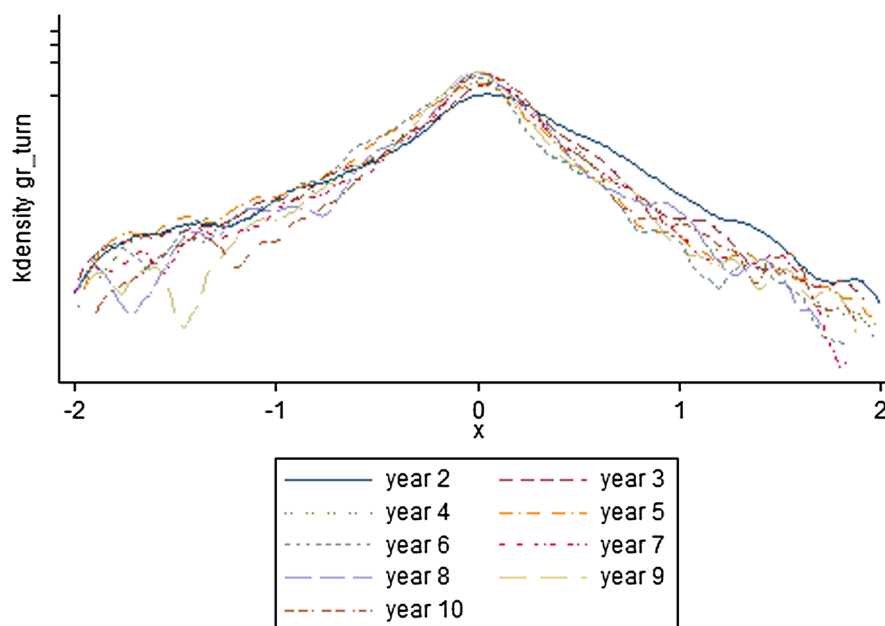
exponential distribution found in other work (Bottazzi and Secchi 2006; Coad and Tamvada 2012; Daunfeldt and Halvarsson 2015). In every year, about half of the firms will have negative growth rates, which emphasises further that our use of the term ‘sales growth’ does not imply that new ventures all have (positive) growth, but that there are many cases of decline (i.e. negative growth rates).

Summary statistics for the explanatory variables are presented in Table 3 in ‘Appendix’.

## 6 Testing the hypotheses

This section presents the crux of our empirical contribution, which can be found in our plots of the evolution of the  $R^2$  statistic over time (see Figs. 3a, b, 4). We present the evolution of the Nagelkerke  $R^2$  statistic for four regression specifications—in some cases we include lagged growth as an explanatory variable (at the cost of losing an extra year’s results), and in some cases we focus on a subsample of relatively large firms (i.e. those with above-median

**Fig. 2** Growth rate distributions for different years. Note the log scale on the y-axis



sales in year 1). Regression results tables for the baseline specification are also presented for the sake of completeness as Tables 4, 5 and 6 in ‘Appendix’.

## 6.1 Plotting the $R^2$ statistics

### 6.1.1 Sales growth

Figure 3a shows how the Nagelkerke  $R^2$  statistic for sales growth regressions evolves over the first 10 years. It starts off in year 2 at values of 28–37 % (depending on the regression specification), which is considerably higher than normally found in the literature on growth rate regressions (no doubt due to our unusually rich information on business behaviour). A closer look at the regression coefficients, reported in Table 4 in ‘Appendix’ for the baseline model, shows that the most significant variables are the bank account activity variables (volatility and overdraft behaviour).

Figure 3a shows that the  $R^2$  decreases in the years after entry for the four specifications shown. In year 2 it is in the range of 27–37 %, whereas by year 10 it is in the range of 13–22 %. Year 5, which corresponds to the deep recession of 2009, does not stand out or interrupt the overall trend. As new ventures age, it seems to become increasingly difficult to accurately predict their growth. This implies the fog seems to thicken and is in line with Lotti et al. (2009), who observe that Gibrat’s Law appears to hold as a ‘long-

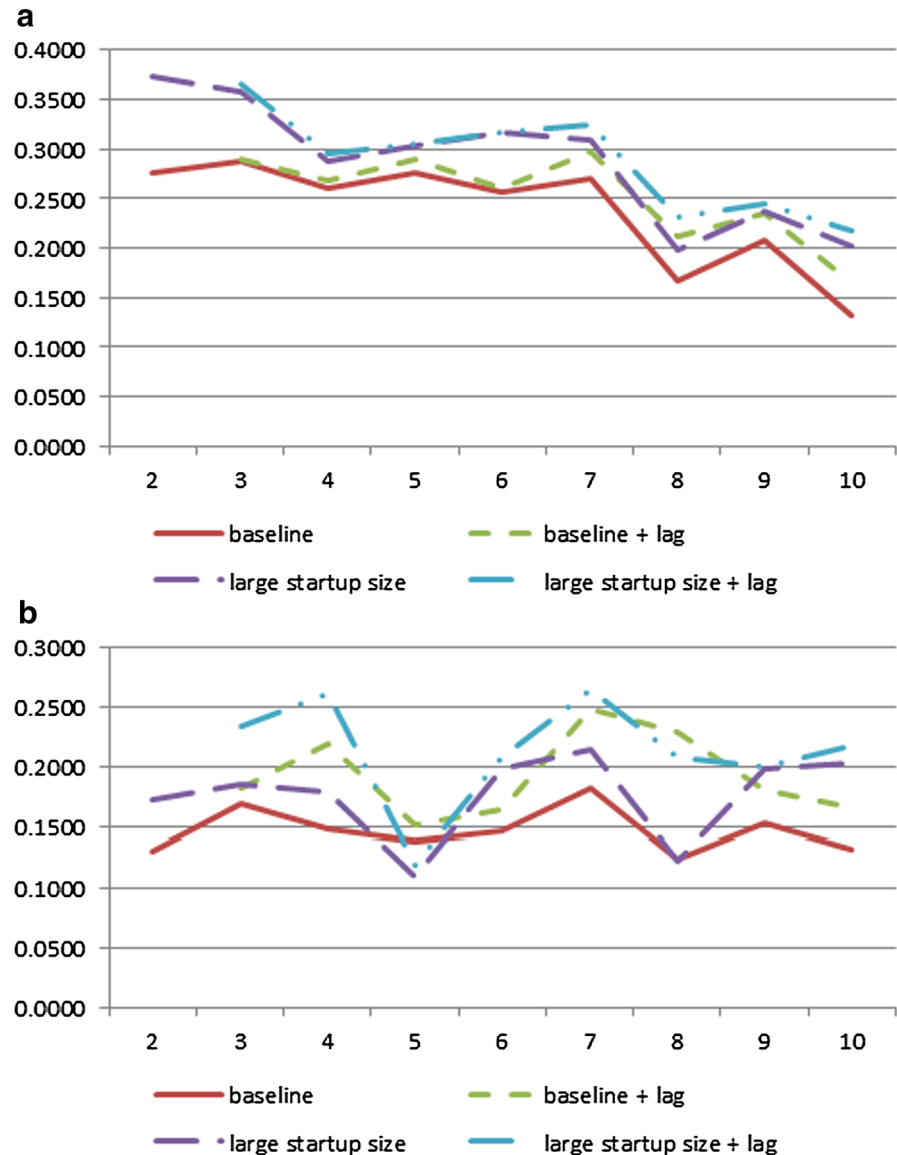
run regularity’ as time goes by, and growth becomes harder to predict.

Another observation is that, from both Fig. 3a, b, for nearly all years, the Nagelkerke  $R^2$  values for the equations that only include the larger enterprises (i.e. the ‘large startup size’ subsample) are higher than for the baseline sample. This implies it is harder to explain the growth performance of smaller firms, which exhibit greater volatility. This may explain, at least in part, why analyses using relatively large and well-established ‘new ventures’ (Hmieleski and Baron 2009; Dencker et al. 2009; Baum and Bird 2010) are able to show higher explanatory power than those included here.

Table 4 in ‘Appendix’ shows this decreasing trend in the explanatory power of our regressions is observed for alternative indicators of goodness of fit—the standard  $R^2$  statistic as well as the Cox–Snell  $R^2$  statistic—for the baseline case. Further explorations show that this is also the case for the 3 other regression specifications (results available upon request).

Figure 3a shows that the explanatory power of growth rate regressions decreases over time, but it does not explain why. Changes in the ‘fog’—our ability to predict growth—could be due to internal developmental factors in new firms, or they could reflect selection effects—whereby the composition of the sample of survivors is affected by the selective exit of certain types of firms.

**Fig. 3** **a** OLS growth regression Nagelkerke  $R^2$  statistics for individual cross sections for the first 10 years, for 4 different growth rate regression specifications and **b** OLS growth regression Nagelkerke  $R^2$  statistics for the first 10 years, for 4 different growth rate regressions (NVs that survive until the end of year 10). Key: Baseline: full sample. Baseline + lag: full sample controlling for lagged growth. Largest startup size: above-median start-up size only. Largest startup size + lag: above-median start-up size subsample, controlling for lagged growth



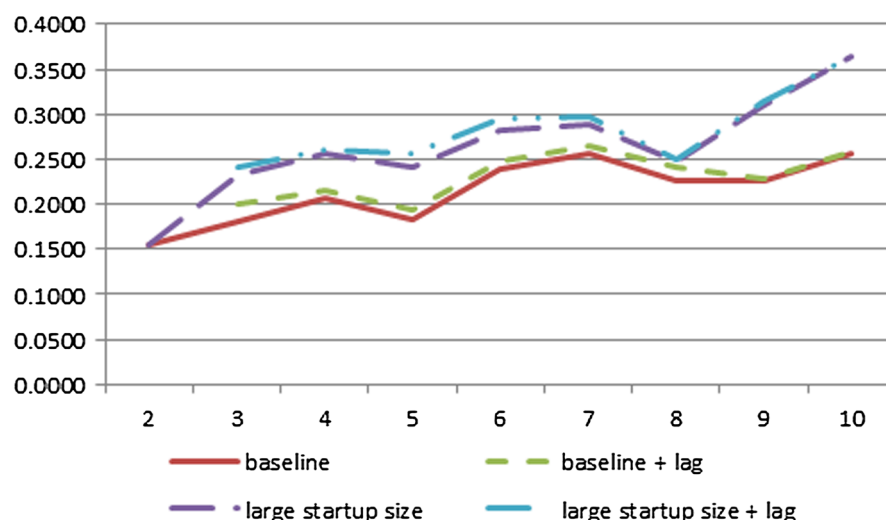
Indeed, previous research has shown that many firms will exit in the years after entry (Audretsch et al. 1999; Santarelli and Vivarelli 2007), and it could be that changes in our ability to explain growth are due to changes in the sample composition over time. One way of eliminating the role of selection effects is to restrict the analysis to only those firms that survive the full 10-year period. Any change in the ability to explain growth for this subsample would then be due to internal developmental factors rather than selection effects.

The results are plotted in Fig. 3b (and the regression results for the baseline case are presented in Table 5 in ‘Appendix’). For the subsample of

surviving firms, the  $R^2$  shows no clear trend over time. For surviving firms, there is no clear change in our ability to explain their growth in the years after entry. Any deterioration in our ability to explain growth in the years after entry (shown in Fig. 3a) would therefore seem to be driven by the relative ease of explaining the growth (or perhaps more precisely: the decline) of short-lived firms.<sup>22</sup>

<sup>22</sup> More specifically, it appears that our ability to explain the relative growth performance of short-lived firms is largely due to the bank account activity variables. When these variables are omitted, we can no longer observe that our ability to explain growth decreases in the post-entry years. (If anything, it seems

**Fig. 4** Logit survival regression: Nagelkerke  $R^2$  statistics for individual cross sections for first 10 years, for 4 different survival regression specifications. Key to regression specifications: the baseline model refers to the full sample with or without controlling for lagged growth. Regressions labelled ‘large startup size’ refer to a subsample of firms with above-median start-up size (i.e. above-median values of sales in the first year)



Overall, therefore, the evidence in Fig. 3a, b suggests that our ability to explain growth deteriorates in the years after entry, with this becoming closer to random over time. This seems to be driven by the changing composition of the sample of surviving firms (i.e. selection effects) rather than any internal developmental factors within firms. Focusing on a core subsample of NVs that survive until the end of year 10 (and thus removing any sample composition effects because we have the same number of observations in each year), our ability to explain growth remains roughly constant over time (Fig. 3b). Overall, this mixed evidence is in keeping with Hypothesis 1.

### 6.1.2 Survival

To test Hypothesis 2 we run year-by-year regressions (presented in detail in Table 6 in ‘Appendix’ for the baseline case) and plot the evolution of the Nagelkerke  $R^2$  statistics in Fig. 4. The ‘fog’ regarding survival—i.e. our ability to explain the survival of firms—seems to clear in the years after entry. Figure 4 shows how the Nagelkerke  $R^2$  starts off at around 15 % in year 2 and increases to 26–36 % by year 10. This is consistent with our simulation model and the predictions of Hypothesis 2.

The key difference between the growth rate regressions (Fig. 3a, b), on the one hand, and the

survival regressions (Fig. 4), on the other hand, is that ‘the fog clears’ in the years after entry when the task is to explain survival, yet it remains dense when the task is to explain growth.

### 6.2 Robustness analysis

Further evidence on the robustness of our findings comes from considering alternative measures of goodness of fit, in addition to the Nagelkerke  $R^2$ . These are shown at the bottom of Tables 4, 5 and 6 in ‘Appendix’. For the growth regressions, the  $R^2$  and Cox–Snell  $R^2$  statistics closely mirror the Nagelkerke  $R^2$ , with no clear trend in the  $R^2$  statistic. For survival, we report the Cox–Snell  $R^2$ , as well as information on the percentage of cases correctly classified: the latter increase in most years after start-up, hence confirming our earlier results using the Nagelkerke  $R^2$  statistic. The Cox–Snell measure provides results that are less clear-cut, however: although it rises during the early years, it reaches a peak in year 6.

Another way of exploring the robustness of our results is by taking an alternative regression specification with a different set of explanatory variables. Earlier we commented on the fact that our database contains a number of variables relating to bank account activity, which constitute a rich and unique source of information on firm behaviour, although these variables remain little-known in the literature, and also they may raise concerns of endogeneity (e.g. risky unauthorised overdraft behaviour may be a cause or a consequence of poor performance in terms of growth or survival

Footnote 22 continued

that our ability increases post-entry, when bank account activity variables are not included.) These extra results are available from the corresponding author upon request.

prospects). We repeated the analysis excluding the variables relating to bank account volatility and obtained the following results. For the growth rate regressions, a first observation was that the Nagelkerke  $R^2$  statistics were very low, in the range of 3–7 % for our baseline specification. If anything, the Nagelkerke  $R^2$  statistics appeared to *increase* slightly in the years after entry, although this increase was not monotonic. When the growth rate regressions were performed on the core sub-sample of firms surviving until the end of the 10-year period, the Nagelkerke  $R^2$  generally decreased, if anything, in the years after entry. Our clearest results were observed in our survival regressions, where the Nagelkerke  $R^2$  followed an increasing trend in the years after entry. All in all, when we repeated the analysis without the bank account activity variables, our results for survival were relatively clear in showing that the survival ‘fog’ tends to clear in the years after entry (i.e. that the Nagelkerke  $R^2$  generally increases in the years after entry). Our results for growth were less clear-cut, probably because after dropping the bank account variables the overall explanatory power was very low (Nagelkerke  $R^2$  statistics of around 5 % or lower) and hence the lower signal-to-noise ratio made it hard to detect any clear trend.

## 7 Conclusion

Business owners, providers of finance, and governments have much to gain from developing a better understanding of the factors influencing the performance of new ventures (NVs) in the years after entry. The starting point for this paper was that the post-entry performance of new ventures is highly diverse, the selection environment is noisy (characterised by imperfect mechanisms of survival/growth of the fittest), and that our ability to explain and perhaps forecast the survival/growth in new ventures is weak. In the terminology of this paper, the fog was thick. The challenge therefore was to examine whether, as the new venture aged, it became easier to explain performance: did the fog lift? Our final question was, if the fog does lift with time, did visibility improve in steps or stages (Phelps et al. 2007; Levie and Lichtenstein 2010) or was the process more continuous?

To address these questions we primarily drew upon a theoretical framework that sees new venture sales growth as a random walk (Levinthal 1991; Le Mens et al. 2011), and survival being determined by

the stock of available resources (proxied by size), where these resources are either present at start-up or accumulated after entry. We used this theory to derive testable hypotheses that our ability to explain growth (i.e. the  $R^2$  from growth regressions) should remain low over time, but that our ability to explain survival should increase in the years after entry.

We conducted our tests on 6579 new ventures which, because they were genuinely representative of NVs, were on balance considerably smaller than those identified in prior work. These NVs were tracked over the years 2004–2014, generating two key findings. First, in the sales growth regressions, the goodness-of-fit measure (Nagelkerke  $R^2$ ) decreases in value in the years after entry—implying that our ability to explain firm growth deteriorates, or that ‘the fog thickens’. However, when we sidestep issues of ‘selection’ and focus only on a subsample of NVs that we know will survive until the end of the period of observation, then our ability to explain the growth in this subsample of survivors remains low but does not change over time. Hence, any decrease in our ability to explain growth in the years after entry appears to be driven by the presence of short-lived firms, rather than being due to internal developmental factors within surviving firms. In any case, our ability to explain growth remains low throughout the period investigated.<sup>23</sup> Second, in the survival equations, using three performance metrics we find that, on balance, the goodness-of-fit increases in years since start-up. This suggests that the fog does lift somewhat with time when the task is to predict survival.

In terms of the questions posed at the start of the paper, we take our evidence as showing that the growth rate fog is always thick and shows no signs of improvement with time, in line with our theory. Survival visibility, however, does seem to improve with time, but not in a clear ‘step’ fashion.

Finally, we see important areas for developing this approach. Currently our data track a cohort of new

<sup>23</sup> Our finding that there is little predictability in sales growth is in itself a valuable finding, with many implications for scholars and providers of finance. Some (admittedly speculative) possible implications can be mentioned here. First, stakeholders might be encouraged to take a broad-based approach to investing, rather than trying to invest in just one firm suspected of soon embarking on fast growth; and second, it hints that there might be little to be gained for investors from investing too heavily in collecting detailed information on firms to predict their growth performance—because growth is so hard to predict.



ventures during an unusual period—beginning in benign macro-economic conditions that are followed by a deep recession. Ideally we would like to know whether our findings hold under different macro-conditions. However, future efforts in this direction will face challenges of obtaining comprehensive datasets on NVs (from year 1) that also include a rich set of explanatory variables.

**Acknowledgments** We are grateful to Jose Garcia Quevedo, Gabriele Pellegrino, Maria Savona, Karl Wennberg, and participants at the RATIO Institute Stockholm and DRUID 2013 (ESADE, Barcelona) for many helpful comments. A.C. gratefully acknowledges financial support from the ESRC, TSB, BIS and NESTA on grants ES/H008705/1 and ES/J008427/1 as part of the IRC distributed projects initiative, as well as from the AHRC as part of the FUSE project. The views expressed are purely those of the authors and may not in any circumstances be

regarded as stating an official position of the European Commission. J.S.F. and R.G.R. write only in a personal capacity and do not necessarily reflect the views of Barclays Bank. Any remaining errors are ours alone.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix

See Tables 2, 3, 4, 5 and 6.

**Table 2** Variable names and definitions

	Description	Prior work	Relationship with growth and survival
<b>Dependent variables</b>			
Open	=1 if enterprise continues to trade at end of period (Open = 0 if the enterprise exited)		
Growth rate	Growth is measured as Credit turnover—value of payments into a current account—excluding payments from a related account (deposit account). In our baseline estimates, the annual growth rate is the log-differences of turnover ( $\log(\text{turnover}(t)) - \log(\text{turnover}(t-1))$ ), which is then normalised with mean = 0, SD = 1		
<b>Independent variables</b>			
Lagged (log) turnover	Credit turnover—the value of payments into a current account, excluding payments from related accounts, e.g. deposit accounts held by the business—is the metric of size. It is a very close approximation to sales revenue (turnover) inclusive of taxes	Coad et al. (2013)	+ with survival
<b>Owner(s) characteristics</b>			
Age	Mean age of owner(s) at start-up	Persson (2004)	+ with survival
Age squared	Square of age	Persson (2004)	+ with survival – with growth
Education dummies	Highest level of educational attainment by owner(s): 1 = none; 2 = GCSE; 3 = A-level; 4 = Degree or higher	Dencker et al. (2009) Parker (2009)	+ with survival Mixed with growth and survival
Business experience: family	business experience of parents: 0 = No; 1 = Yes	Gimeno et al. (1997)	+ with Survival n/s with growth
Business experience: self	Previous business experience of Owner: 0 = No; 1 = Yes	Dencker et al. (2009)	+ with survival when combined with other variables
No owners	Owners in excess of minimum number for legal form: 0 = No; 1 = Yes	Klotz et al. (2014)	Mixed results
Male owner(s)	At least one male owner: 0 = No; 1 = Yes	Gilbert et al. (2006), Klotz et al. (2014) Dencker et al., (2009)	Mixed results with both growth and survival No gender effect
Sources of advice	See individual sources identified below	Solomon et al. (2013), Chrisman and McMullan (2004) Pons Rotger et al. (2012)	+ with survival and growth + (small) with survival and growth

**Table 2** continued

	Description	Prior work	Relationship with growth and survival
Enterprise agency/business link EABL	Advice/support (prior to start-up) from enterprise agency/business link: 0 = No; 1 = Yes	Mole et al. (2011)	+ with growth in some cases
Accountant	Advice/support (prior to start-up) from Accountant: 0 = No; 1 = Yes	Frankish et al. (2011)	n/s with growth or survival
Solicitor	Advice/support (prior to start-up) from Solicitor: 0 = No; 1 = Yes	Frankish et al. (2011)	n/s with growth or survival
College	Advice/support (prior to start-up) from a College: 0 = No; 1 = Yes	Frankish et al. (2011)	n/s with growth or survival
Start-right seminar	Participation in Barclays Start-right seminar: 0 = No; 1 = Yes	Coad et al. (2016)	n/s with survival and growth + with bank loyalty
Princes Youth Business Trust	Advice/support (prior to start-up) from Princes Youth Business Trust : 0 = No; 1 = Yes	Greene (2009)	n/s using sophisticated evaluation methods
Family	Advice/support (prior to start-up) from Family: 0 = No; 1 = Yes	Frankish et al. (2011)	n/s with growth or survival
Other	Advice/support (prior to start-up) from any other source: 0 = No; 1 = Yes	Frankish et al. (2011)	n/s with growth or survival
Bank account activity variables			
Volatility	Ratio of the standard deviation of monthly turnover to the mean monthly turnover, summed over two six-month periods to obtain annual volatility indicator	Coad et al. (2013)	– with survival – with growth
Overdraft excess	=1 if in excess of authorised overdraft limit at any time	Coad et al. (2013), Frankish et al. (2012)	– with survival n/s with growth
Overdraft excess duration	Proportion of period in excess of authorised overdraft limit	Coad et al. (2013), Frankish et al. (2012)	– with survival – with growth
Authorised overdraft use	=1 if authorised overdraft used at any time	Coad et al. (2013), Frankish et al. (2012)	n/s with survival + with growth
Extent of auth. OD use	Mean proportion of authorised overdraft limit used	Coad et al. (2013), Frankish et al. (2012)	n/s with survival – with growth
Legal form dummies			
Legal form	1 = Company, 2 = Partnership, 3 = Sole Trader	Storey (1994)	Limited liability companies + with growth and survival
Industry dummies	Business activity: 1 = Agriculture, 2 = Manufacturing, 3 = Construction, 4 = Motor Trades, 5 = Wholesale, 6 = Retail, 7 = Hotels & Catering, 8 = Transport, 9 = Property Services, 10 = Business Services, 11 = Health, Education and Social Work, 12 = Other Services	Dencker et al. (2009)	Dencker et al. (2009) has three sectors. n/s with survival
Region dummies	Region: 1 = East Midlands, 2 = East of England, 3 = London, 4 = North East, 5 = North West, 6 = South East, 7 = South West, 8 = West Midlands, 9 = Yorkshire and The Humber, 10 = Wales	Botham and Graves (2011)	+ London and South-East for Growth n/s for survival

**Table 3** Summary statistics for the explanatory variables, for year 1

Variable	Mean	SD	Min	Max	Obs.
Age	39.036	10.225	16.2	78.16	6570
Education dummies					
<NVQ2 (none)	0.228	0.419	0	1	6579
NVQ2 (GCSE)	0.332	0.471	0	1	6579
NVQ3 (A-level)	0.171	0.377	0	1	6579
NVQ4 (degree or higher)	0.270	0.444	0	1	6579

**Table 3** continued

Variable	Mean	SD	Min	Max	Obs.
Family business experience	0.633	0.482	0	1	6579
Personal business experience	0.720	0.449	0	1	6579
Sources of advice					
EABL	0.102	0.303	0	1	6579
Accountant	0.361	0.480	0	1	6579
Solicitor	0.049	0.215	0	1	6579
College	0.040	0.196	0	1	6579
Start-right seminar	0.007	0.086	0	1	6579
Prince's youth business trust	0.014	0.116	0	1	6579
Family	0.300	0.458	0	1	6579
Other	0.064	0.245	0	1	6579
Bank account activity					
Volatility	0.844	0.603	0	4.843	5524
Auth. OD use	0.186	0.389	0	1	5524
Extent of auth. OD use	0.041	0.118	0	0.93	5524
Unauthorised OD use	0.407	0.491	0	1	5524
Extent of unauthorised OD use	0.042	0.103	0	1	5524
Excess owners	0.156	0.363	0	1	6562
Male owners	0.963	0.631	0	5	6562
Legal form dummies					
Company	0.372	0.483	0	1	6579
Partnership	0.133	0.340	0	1	6579
Sole trader	0.495	0.500	0	1	6579
Industry dummies					
Agriculture	0.010	0.100	0	1	6579
Manufacturing	0.049	0.215	0	1	6579
Construction	0.147	0.354	0	1	6579
Retail	0.178	0.382	0	1	6579
Transport	0.027	0.161	0	1	6579
Accommodation	0.094	0.292	0	1	6579
Information	0.060	0.237	0	1	6579
Real estate	0.035	0.183	0	1	6579
Professional	0.072	0.258	0	1	6579
Administrative	0.150	0.357	0	1	6579
Education	0.008	0.090	0	1	6579
Health	0.016	0.125	0	1	6579
Arts	0.034	0.182	0	1	6579
Other	0.122	0.327	0	1	6579
Region dummies					
East Midlands	0.073	0.261	0	1	6570
East of England	0.155	0.362	0	1	6570
London	0.223	0.416	0	1	6570
North East	0.037	0.189	0	1	6570
North West	0.067	0.251	0	1	6570
South East	0.127	0.333	0	1	6570
South West	0.098	0.298	0	1	6570
West Midlands	0.095	0.294	0	1	6570
Yorkshire and Humber	0.061	0.239	0	1	6570
Wales	0.063	0.242	0	1	6570

**Table 4** Growth rate regressions for each cross section, for the full sample of firms (without controlling for lagged growth)

	(1) Year 2	(2) Year 3	(3) Year 4	(4) Year 5	(5) Year 6	(6) Year 7	(7) Year 8	(8) Year 9	(9) Year 10
Lagged (log) turnover	-0.180*** (0.0166)	-0.191*** (0.0161)	-0.169*** (0.0179)	-0.237*** (0.0341)	-0.150*** (0.0202)	-0.163*** (0.0225)	-0.133*** (0.0240)	-0.123*** (0.0236)	-0.101*** (0.0222)
Age	-0.0214** (0.00964)	0.0116 (0.00927)	0.00809 (0.00995)	0.00726 (0.0110)	0.00335 (0.0109)	0.00718 (0.0116)	-0.00312 (0.0122)	0.00177 (0.0133)	0.00565 (0.0137)
Age squared	0.000195* (0.000110)	-0.000157 (0.000108)	-9.77e-05 (0.000115)	-0.000104 (0.000129)	-4.99e-05 (0.000127)	-8.55e-05 (0.000134)	3.10e-05 (0.000141)	-3.88e-05 (0.000151)	-8.45e-05 (0.000156)
Education dummies (omitted = none)									
GCSE	-0.0118 (0.0324)	-0.0889** (0.0377)	0.00798 (0.0393)	0.0545 (0.0445)	0.00255 (0.0451)	0.0245 (0.0481)	0.0169 (0.0450)	0.0335 (0.0500)	-0.0574 (0.0539)
A-level	-0.0264 (0.0401)	-0.0322 (0.0435)	0.0204 (0.0482)	0.0543 (0.0556)	-0.126** (0.0555)	0.0935 (0.0618)	0.0585 (0.0592)	0.0137 (0.0656)	-0.0174 (0.0664)
Degree or higher	0.00186 (0.0391)	-0.0496 (0.0468)	0.00257 (0.0474)	0.160*** (0.0528)	0.0155 (0.0539)	0.132** (0.0568)	0.0620 (0.0561)	0.104 (0.0668)	0.0803 (0.0624)
Business experience									
Family	0.00198 (0.0263)	0.00955 (0.0290)	-0.0598* (0.0316)	0.113*** (0.0347)	0.0570 (0.0368)	0.0530 (0.0360)	0.0409 (0.0364)	-0.0208 (0.0394)	0.0464 (0.0394)
Self	0.000326 (0.0308)	0.0137 (0.0328)	-0.00835 (0.0358)	0.0551 (0.0422)	0.0371 (0.0404)	0.00632 (0.0409)	0.00454 (0.0422)	0.0447 (0.0478)	-0.0225 (0.0482)
Sources of advice									
Enterprise agency/business link (EABL)	-0.0427 (0.0439)	0.0700 (0.0556)	-0.0324 (0.0613)	0.00278 (0.0612)	0.0121 (0.0582)	-0.111* (0.0652)	-0.0239 (0.0738)	0.0363 (0.0515)	0.0210 (0.0594)
Accountant	-0.00961 (0.0272)	-0.00276 (0.0300)	-0.0476 (0.0330)	0.0240 (0.0388)	-0.0305 (0.0375)	-0.0524 (0.0367)	-0.00756 (0.0379)	0.0269 (0.0400)	0.0954** (0.0401)
Solicitor	0.0180 (0.0802)	-0.0537 (0.0753)	0.0888 (0.0717)	-0.0935 (0.0999)	0.152 (0.0962)	-0.0365 (0.0881)	-0.0194 (0.0883)	-0.0502 (0.107)	-0.0162 (0.0978)
College	0.00454 (0.0633)	0.0600 (0.0613)	0.0680 (0.0517)	0.0248 (0.0718)	0.0679 (0.0721)	-0.0595 (0.0777)	0.0272 (0.110)	0.00801 (0.104)	-0.145 (0.103)
Start-right seminar	-0.123 (0.177)	0.0806 (0.194)	-0.124 (0.233)	-0.149 (0.122)	0.406* (0.233)	0.363* (0.197)	0.337* (0.186)	0.482*** (0.161)	0.168 (0.116)
Princes Youth Business Trust	-0.00364 (0.141)	-0.203 (0.169)	0.0349 (0.135)	-0.531* (0.280)	0.340 (0.216)	-0.0896 (0.284)	0.335 (0.304)	0.0680 (0.310)	0.175 (0.237)
Family	0.0289 (0.0290)	0.0450 (0.0339)	0.00718 (0.0334)	-0.0615 (0.0395)	-0.104** (0.0423)	-0.00892 (0.0424)	0.0350 (0.0470)	-0.0479 (0.0460)	-0.00179 (0.0452)

Table 4 continued

	(1) Year 2	(2) Year 3	(3) Year 4	(4) Year 5	(5) Year 6	(6) Year 7	(7) Year 8	(8) Year 9	(9) Year 10
Other	0.0579 (0.0584)	-0.0515 (0.0642)	-0.0923 (0.0597)	0.102 (0.0807)	-0.122 (0.0886)	0.146** (0.0693)	-0.0547 (0.0875)	-0.109 (0.112)	-0.0443 (0.0701)
Bank account activity variables									
Volatility	-0.639*** (0.0374)	-0.626*** (0.0478)	-0.561*** (0.0521)	-0.606*** (0.0583)	-0.508*** (0.0736)	-0.486*** (0.0662)	-0.358*** (0.0692)	-0.356*** (0.0728)	-0.287*** (0.0713)
Overdraft excess	0.0518* (0.0301)	0.0106 (0.0343)	0.0715** (0.0348)	-0.0202 (0.0426)	-0.0104 (0.0441)	0.0174 (0.0468)	0.00951 (0.0497)	0.0268 (0.0553)	0.0472 (0.0535)
Overdraft excess duration	-0.971*** (0.146)	-0.985*** (0.142)	-1.024*** (0.157)	-0.739*** (0.168)	-1.017*** (0.242)	-1.189*** (0.269)	-0.677*** (0.262)	-1.385*** (0.406)	-0.837** (0.440)
Authorised overdraft use	0.279*** (0.0414)	0.152*** (0.0387)	0.00449 (0.0480)	0.0956* (0.0524)	0.0455 (0.0736)	0.152*** (0.0536)	0.0729 (0.0491)	0.0994* (0.0528)	-0.0561 (0.0668)
Extent of auth. OD use	-0.257*** (0.0791)	-0.173** (0.0754)	-0.0335 (0.0785)	-0.0774 (0.0906)	-0.133 (0.129)	-0.158 (0.107)	-0.0323 (0.121)	-0.238*** (0.0912)	0.113 (0.129)
No. owners	0.101** (0.0427)	0.113** (0.0464)	0.0208 (0.0474)	0.0335 (0.0536)	0.0489 (0.0490)	-0.0387 (0.0492)	-0.0278 (0.0485)	-0.0238 (0.0499)	0.0471 (0.0512)
Male owner(s)	0.0957*** (0.0337)	0.0506 (0.0428)	0.0768* (0.0460)	0.0578 (0.0505)	-0.0577 (0.0464)	0.0665 (0.0579)	0.142** (0.0687)	-0.0204 (0.0588)	0.0426 (0.0682)
Legal form dummies (omitted = Company)									
Partnership	-0.0701* (0.0391)	-0.124*** (0.0461)	-0.129*** (0.0493)	-0.240*** (0.0647)	-0.173*** (0.0665)	-0.211*** (0.0641)	-0.101* (0.0546)	-0.108* (0.0619)	-0.0233 (0.0603)
Sole trader	-0.134*** (0.0359)	-0.189*** (0.0383)	-0.229*** (0.0434)	-0.196*** (0.0587)	-0.0772 (0.0499)	-0.159*** (0.0497)	-0.173*** (0.0485)	-0.184*** (0.0592)	0.0118 (0.0512)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.748*** (0.298)	2.255*** (0.284)	2.122*** (0.322)	2.887*** (0.470)	1.954*** (0.384)	1.934*** (0.392)	1.659*** (0.399)	1.744*** (0.412)	1.358*** (0.426)
Observations	4147	3201	2584	2146	1819	1600	1421	1308	1205
R-squared	0.258	0.268	0.240	0.257	0.235	0.244	0.147	0.185	0.116
LL	-5013	-3861	-2928	-2513	-2005	-1634	-1394	-1312	-1169
LL(0)	-5630	-4360	-3283	-2832	-2248	-1858	-1507	-1446	-1243
Cox-Snell $R^2$	0.258	0.268	0.240	0.257	0.235	0.244	0.147	0.185	0.116
Nagelkerke $R^2$	0.276	0.287	0.261	0.277	0.257	0.271	0.167	0.208	0.132

Robust standard errors in parentheses

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

**Table 5** Growth rate regressions for each cross section, for firms that survive until the end of year 10 (without controlling for lagged growth)

	(1) Year 2	(2) Year 3	(3) Year 4	(4) Year 5	(5) Year 6	(6) Year 7	(7) Year 8	(8) Year 9	(9) Year 10
Lagged (log) turnover	-0.117*** (0.0252)	-0.122*** (0.0219)	-0.117*** (0.0239)	-0.097*** (0.0178)	-0.0695*** (0.0165)	-0.105*** (0.0235)	-0.115*** (0.0264)	-0.124*** (0.0256)	-0.101*** (0.0222)
Age	-0.0367*** (0.0108)	-0.00495 (0.00994)	-0.00577 (0.00999)	0.00343 (0.00957)	-0.0193* (0.0102)	0.0123 (0.0117)	0.000883 (0.0132)	-0.00720 (0.0137)	0.00565 (0.0137)
Age squared	0.000353*** (0.000121)	5.81e-05 (0.000113)	2.91e-05 (0.000114)	-4.99e-05 (0.000113)	0.000198* (0.000115)	-0.000141 (0.000129)	-1.77e-05 (0.000152)	6.45e-05 (0.000157)	-8.45e-05 (0.000156)
Education dummies (omitted = none)									
GCSE	-0.0372 (0.0470)	0.0295 (0.0418)	-0.00948 (0.0373)	0.0200 (0.0387)	-0.0106 (0.0421)	-0.0201 (0.0386)	0.00310 (0.0443)	0.0267 (0.0514)	-0.0574 (0.0539)
A-level	-0.00554 (0.0551)	0.0382 (0.0513)	-0.0103 (0.0528)	0.00809 (0.0494)	-0.0912 (0.0559)	0.0592 (0.0514)	0.0619 (0.0581)	0.0285 (0.0639)	-0.0174 (0.0664)
Degree or higher	-0.0490 (0.0584)	0.0250 (0.0529)	-0.0677 (0.0475)	0.0631 (0.0475)	0.00755 (0.0516)	0.0166 (0.0497)	0.0248 (0.0583)	0.0954 (0.0688)	0.0803 (0.0624)
Business experience									
Family	0.0148 (0.0368)	0.0248 (0.0359)	-0.00885 (0.0332)	0.0235 (0.0321)	-0.0228 (0.0332)	-0.00852 (0.0318)	0.0467 (0.0366)	-0.0115 (0.0397)	0.0464 (0.0394)
Self	0.0233 (0.0457)	0.00953 (0.0381)	0.0365 (0.0406)	0.0263 (0.0358)	0.0475 (0.0377)	-0.00373 (0.0405)	0.00269 (0.0457)	0.0402 (0.0511)	-0.0225 (0.0482)
Sources of advice									
Enterprise agency/business link (EABL)	-0.0287 (0.0719)	0.0638 (0.0568)	0.00346 (0.0457)	-0.00304 (0.0558)	0.0660 (0.0578)	-0.0332 (0.0621)	-0.0332 (0.0813)	0.00160 (0.0519)	0.0210 (0.0594)
Accountant	0.000112 (0.0391)	0.000980 (0.0355)	-0.0261 (0.0316)	-0.0345 (0.0337)	0.00750 (0.0351)	-0.00326 (0.0343)	-0.00514 (0.0389)	0.00567 (0.0396)	0.0954*** (0.0401)
Solicitor	0.00102 (0.0720)	-0.269*** (0.127)	-0.0256 (0.0882)	0.0689 (0.0943)	0.126* (0.0687)	-0.0654 (0.0880)	-0.0599 (0.0946)	-0.0893 (0.111)	-0.0162 (0.0978)
College	0.0209 (0.0872)	0.0721 (0.0617)	0.0125 (0.0574)	-0.102 (0.0651)	0.0655 (0.0639)	-0.188*** (0.0835)	0.0699 (0.108)	-0.0295 (0.104)	-0.145 (0.103)
Start-right seminar	-0.320*** (0.0939)	-0.183*** (0.0868)	0.0203 (0.203)	-0.314 (0.354)	0.162*** (0.0766)	0.0424 (0.213)	0.288 (0.249)	0.305*** (0.101)	0.168 (0.116)
Princes Youth Business Trust	0.273 (0.325)	0.510 (0.472)	-0.0217 (0.190)	-0.619 (0.482)	0.136 (0.216)	-0.149 (0.322)	0.259 (0.303)	-0.0146 (0.330)	0.175 (0.237)
Family	0.0778* (0.0429)	0.0331 (0.0415)	-0.0282 (0.0347)	-0.00653 (0.0354)	-0.0612 (0.0375)	-0.00967 (0.0441)	0.0126 (0.0528)	-0.0352 (0.0455)	-0.00179 (0.0452)



Table 5 continued

	(1) Year 2	(2) Year 3	(3) Year 4	(4) Year 5	(5) Year 6	(6) Year 7	(7) Year 8	(8) Year 9	(9) Year 10
Other	0.204** (0.0865)	-0.0227 (0.0836)	-0.158** (0.0789)	0.0495 (0.0755)	-0.0882 (0.0872)	0.157** (0.0704)	-0.0264 (0.0828)	-0.148 (0.124)	-0.0443 (0.0701)
Bank account activity variables									
Volatility	-0.140* (0.0838)	-0.340*** (0.0835)	-0.152** (0.0700)	-0.197*** (0.0663)	-0.243*** (0.0813)	-0.206*** (0.0782)	-0.262*** (0.0853)	-0.317*** (0.0875)	-0.287*** (0.0713)
Overdraft excess	-0.0654 (0.0462)	0.00928 (0.0396)	0.110** (0.0455)	-0.0682 (0.0420)	-0.00233 (0.0529)	0.0354 (0.0462)	0.00146 (0.0530)	0.00322 (0.0575)	0.0472 (0.0535)
Overdraft excess duration	-0.524 (0.424)	-0.766*** (0.289)	-0.831** (0.377)	-0.489** (0.236)	-1.073** (0.469)	-1.136*** (0.360)	-0.546** (0.261)	-0.823* (0.468)	-0.837* (0.440)
Authorised overdraft use	0.181*** (0.0523)	0.0904** (0.0433)	0.0840** (0.0419)	0.0881* (0.0487)	0.0886 (0.0591)	0.103** (0.0462)	0.0812* (0.0477)	0.120** (0.0518)	-0.0561 (0.0668)
Extent of auth. OD use	-0.00699 (0.103)	0.00374 (0.0884)	-0.106 (0.0794)	-0.105 (0.0873)	-0.221** (0.109)	-0.110 (0.117)	-0.0297 (0.128)	-0.228** (0.0970)	0.113 (0.129)
No. owners	0.0959* (0.0539)	0.0538 (0.0508)	0.0573 (0.0434)	0.0157 (0.0440)	0.0241 (0.0437)	-0.0171 (0.0428)	0.0150 (0.0486)	0.0118 (0.0507)	0.0471 (0.0512)
Male owner(s)	-0.0363 (0.0516)	0.0352 (0.0516)	0.00650 (0.0455)	0.0311 (0.0452)	-0.0264 (0.0430)	0.00436 (0.0550)	0.141** (0.0694)	-0.0357 (0.0616)	0.0426 (0.0682)
Legal form dummies (omitted = Company)									
Partnership	-0.0857 (0.0641)	-0.150** (0.0586)	-0.127* (0.0680)	-0.0858* (0.0517)	-0.0427 (0.0565)	-0.206*** (0.0569)	-0.0792 (0.0568)	-0.105* (0.0609)	-0.0233 (0.0603)
Sole trader	-0.157*** (0.0466)	-0.186*** (0.0436)	-0.188*** (0.0477)	-0.0825* (0.0425)	-0.0564 (0.0475)	-0.131*** (0.0489)	-0.127*** (0.0479)	-0.209*** (0.0602)	0.0118 (0.0512)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.368*** (0.385)	1.796*** (0.387)	1.608*** (0.392)	1.185*** (0.330)	1.248*** (0.352)	1.117*** (0.427)	1.173*** (0.414)	1.961*** (0.439)	1.358*** (0.426)
Observations	1205	1205	1205	1205	1205	1205	1205	1205	1205
R-squared	0.112	0.144	0.119	0.112	0.121	0.152	0.106	0.135	0.116
LL	-1091	-1033	-918.0	-920.6	-972.4	-969.1	-1110	-1163	-1169
LL(0)	-1162	-1127	-994.6	-992.3	-1050	-1068	-1178	-1250	-1243
Cox-Snell $R^2$	0.112	0.144	0.119	0.112	0.121	0.152	0.106	0.135	0.116
Nagelkerke $R^2$	0.131	0.170	0.148	0.139	0.147	0.183	0.124	0.154	0.132

Robust standard errors in parentheses

\*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

**Table 6** Survival regressions for each cross section, for the full sample of firms (without controlling for lagged growth)

Variables	(1) Open ( <i>t</i> + 1) Year 2	(2) Open ( <i>t</i> + 1) Year 3	(3) Open ( <i>t</i> + 1) Year 4	(4) Open ( <i>t</i> + 1) Year 5	(5) Open ( <i>t</i> + 1) Year 6	(6) Open ( <i>t</i> + 1) Year 7	(7) Open ( <i>t</i> + 1) Year 8	(8) Open ( <i>t</i> + 1) Year 9	(9) Open ( <i>t</i> + 1) Year 10
Log turnover	0.145*** (0.0294)	0.0745** (0.0324)	0.110*** (0.0384)	0.150*** (0.0454)	0.238*** (0.0530)	0.189*** (0.0607)	0.319*** (0.0689)	0.124 (0.0845)	0.264*** (0.0854)
Age	0.0415* (0.0225)	0.0340 (0.0258)	0.0563* (0.0340)	0.0994*** (0.0371)	0.0760 (0.0481)	0.115** (0.0517)	0.0339 (0.0619)	0.116 (0.0711)	0.0831 (0.0819)
Age squared	-0.000391 (0.000267)	-0.000301 (0.000303)	-0.000537 (0.000405)	-0.000939** (0.000443)	-0.000897 (0.000565)	-0.00128** (0.000586)	-0.000247 (0.000722)	-0.00128 (0.000795)	-0.00110 (0.000956)
Education dummies (omitted = none)									
GCSE	0.0615 (0.0934)	-0.0312 (0.112)	-0.0931 (0.139)	0.177 (0.157)	-0.0131 (0.186)	0.00565 (0.222)	0.0424 (0.238)	-0.0297 (0.283)	0.561* (0.312)
A-level	0.176 (0.112)	-0.0885 (0.128)	-0.0250 (0.166)	-0.0606 (0.184)	-0.0185 (0.231)	-0.135 (0.261)	0.477 (0.300)	0.0725 (0.345)	0.396 (0.369)
Degree or higher	-0.153 (0.102)	0.0166 (0.126)	-0.208 (0.151)	0.0879 (0.174)	-0.152 (0.207)	0.298 (0.258)	0.322 (0.277)	-0.00378 (0.345)	0.513 (0.337)
Business experience									
Family	-0.00363 (0.0721)	-0.0560 (0.0867)	-0.0201 (0.107)	-0.182 (0.122)	0.0934 (0.143)	0.0312 (0.174)	0.248 (0.188)	0.130 (0.222)	0.172 (0.251)
Self	-0.0768 (0.0802)	0.152 (0.0953)	0.0161 (0.122)	-0.226 (0.138)	-0.0494 (0.168)	0.0131 (0.195)	-0.197 (0.205)	-0.668** (0.286)	-0.0285 (0.297)
Sources of advice									
Enterprise agency/ business link (EABL)	-0.143 (0.111)	0.0832 (0.146)	0.0555 (0.175)	0.248 (0.206)	0.244 (0.251)	0.434 (0.296)	-0.248 (0.314)	-0.425 (0.375)	0.998* (0.591)
Accountant	-0.122 (0.0747)	0.0178 (0.0895)	0.0441 (0.110)	-0.0422 (0.126)	0.119 (0.149)	-0.154 (0.178)	-0.164 (0.187)	-0.0117 (0.241)	0.0741 (0.248)
Solicitor	0.0536 (0.170)	-0.190 (0.181)	-0.222 (0.228)	-0.116 (0.287)	0.0641 (0.332)	-0.0252 (0.382)	-0.142 (0.445)	0.994 (0.816)	2.096 (1.371)
College	0.0177 (0.170)	0.405* (0.236)	0.0642 (0.243)	-0.0983 (0.305)	0.358 (0.359)	0.754 (0.594)	0.746 (0.632)	1.094 (0.673)	0.701 (0.713)
Start-right seminar	0.00155 (0.427)	-0.218 (0.581)	1.209 (0.946)	-0.104 (0.711)	-0.215 (0.812)	-1.096* (0.622)	-0.939 (1.398)		-3.688** (1.737)
Princes Youth Business Trust	-0.113 (0.318)	0.777 (0.502)	-0.907** (0.386)	1.136 (0.844)	-0.233 (0.726)	-0.233 (0.960)	0.930 (0.927)		

Table 6 continued

Variables	(1) Open ( $t+1$ ) Year 2	(2) Open ( $t+1$ ) Year 3	(3) Open ( $t+1$ ) Year 4	(4) Open ( $t+1$ ) Year 5	(5) Open ( $t+1$ ) Year 6	(6) Open ( $t+1$ ) Year 7	(7) Open ( $t+1$ ) Year 8	(8) Open ( $t+1$ ) Year 9	(9) Open ( $t+1$ ) Year 10
Family	0.00543 (0.0778)	0.0260 (0.0934)	-0.110 (0.112)	0.314** (0.140)	-0.116 (0.160)	-0.0253 (0.188)	0.149 (0.210)	-0.348 (0.249)	-0.275 (0.272)
Other	-0.0314 (0.139)	-0.0155 (0.167)	-0.328* (0.194)	-0.0638 (0.269)	0.0258 (0.312)	0.473 (0.415)	0.914* (0.495)	-0.668* (0.394)	-0.0917 (0.433)
Bank account activity variables									
Volatility	-0.543*** (0.0638)	-0.725*** (0.0695)	-0.694*** (0.0808)	-0.612*** (0.0932)	-0.779*** (0.109)	-0.794*** (0.127)	-0.608*** (0.149)	-0.894*** (0.172)	-0.623*** (0.174)
Overdraft excess	-0.135* (0.0807)	-0.195** (0.0932)	-0.242** (0.115)	-0.0879 (0.136)	-0.574*** (0.156)	-0.276 (0.199)	-0.905*** (0.206)	-0.171 (0.258)	-0.716*** (0.273)
Overdraft excess duration	-3.666*** (0.412)	-2.520*** (0.315)	-1.861*** (0.287)	-1.335*** (0.357)	-1.496*** (0.407)	-1.664*** (0.463)	0.103 (0.501)	-2.210*** (0.637)	-1.429*** (0.692)
Authorised overdraft use	0.0710 (0.127)	0.225 (0.143)	-0.0647 (0.166)	0.127 (0.197)	-0.350 (0.222)	-0.234 (0.276)	0.362 (0.347)	-0.591* (0.323)	0.359 (0.438)
Extent of auth. OD use	0.0299 (0.417)	-0.195 (0.315)	-0.235 (0.355)	0.275 (0.405)	-0.0538 (0.434)	0.175 (0.538)	-0.736 (0.653)	0.745 (0.631)	-0.682 (0.842)
No. owners	0.181 (0.116)	-0.00809 (0.127)	0.203 (0.158)	0.426** (0.187)	-0.0168 (0.191)	-0.441* (0.248)	-0.0962 (0.257)	0.279 (0.355)	-0.197 (0.325)
Male owner(s)	0.134 (0.0896)	0.163 (0.109)	0.0222 (0.136)	0.413*** (0.153)	0.0851 (0.201)	0.739*** (0.225)	0.480* (0.253)	0.0466 (0.312)	0.608* (0.311)
Legal form dummies (omitted = Company)									
Partnership	-0.396*** (0.115)	-0.613*** (0.132)	-0.571*** (0.163)	-0.157 (0.197)	0.233 (0.244)	-0.0815 (0.319)	0.0298 (0.319)	-0.202 (0.402)	-0.0739 (0.361)
Sole trader	-0.234** (0.0942)	-0.191* (0.110)	-0.0942 (0.136)	-0.0158 (0.154)	0.647*** (0.188)	-0.0126 (0.215)	0.485** (0.234)	-0.265 (0.293)	0.400 (0.298)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.147 (0.686)	0.423 (0.783)	-0.00254 (0.923)	-2.351** (1.078)	-2.194* (1.321)	12.59*** (1.535)	-2.069 (1.767)	1.271 (2.167)	10.74*** (2.179)
Observations	5499	4147	3201	2557	2146	1819	1562	1393	1302
LL	-2768	-1965	-1348	-1022	-757.5	-538.2	-460.0	-320.5	-285.3

Table 6 continued

Variables	(1) Open ( <i>t</i> + 1) Year 2	(2) Open ( <i>t</i> + 1) Year 3	(3) Open ( <i>t</i> + 1) Year 4	(4) Open ( <i>t</i> + 1) Year 5	(5) Open ( <i>t</i> + 1) Year 6	(6) Open ( <i>t</i> + 1) Year 7	(7) Open ( <i>t</i> + 1) Year 8	(8) Open ( <i>t</i> + 1) Year 9	(9) Open ( <i>t</i> + 1) Year 10
LL(0)	−3067	−2227	−1569	−1171	−915.9	−668.9	−556.1	−392.1	−360.1
Cox–Snell $R^2$	0.103	0.119	0.129	0.110	0.137	0.134	0.116	0.0977	0.109
Nagelkerke $R^2$	0.154	0.180	0.207	0.184	0.239	0.257	0.227	0.227	0.256
% Correctly classified	76.63	78.85	82.04	83.38	85.97	88.62	89.05	92.75	92.47

Robust standard errors in parentheses

 \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ 

## References

- Anyadike-Danes, M., Bjuggren, C.-M., Gottschalk, S., Hölzl, W., Johansson, D., Maliranta, M., & Myrann, A. (2015). An international cohort comparison of size effects on job growth. *Small Business Economics*, 44, 821–844. doi:10.1007/s11187-014-9622-0.
- Anyadike-Danes, M., & Hart, M., (2014). All grown up? The fate after 15 years of the quarter of a million UK firms born in 1998. Sept 14th, Mimeo.
- Audretsch, D. B., Santarelli, E., & Vivarelli, M. (1999). Start-up size and industrial dynamics: Some evidence from Italian Manufacturing. *International Journal of Industrial Organization*, 17, 965–983.
- Baum, J. R., & Bird, B. J. (2010). The successful intelligence of high growth entrepreneurs: Links to new venture growth. *Organization Science*, 21(2), 397–412.
- Bertrand, M., & Schoar, A. (2003). *Managing with style: The effect of managers on firm policies*. MIT Sloane School of Management, Working Paper 4802-02.
- Botham, R., & Graves, A. (2011). Regional variations in new firm job creation: The contribution of high growth start-ups. *Local Economy*, 26(2), 95–107.
- Bottazzi, G., & Secchi, A. (2006). Explaining the distribution of firm growth rates. *Rand Journal of Economics*, 37(2), 235–256.
- Chrisman, J. J., & McMullan, W. E. (2004). Outsider assistance as a knowledge resource for new venture survival. *Journal of Small Business Management*, 42(3), 229–244.
- Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence*. Cheltenham: Edward Elgar.
- Coad, A. (2014). Death is not a success: Reflections on business exit. *International Small Business Journal*, 32(7), 721–732.
- Coad, A., Frankish, J. S., Nightingale, P., & Roberts, R. G. (2014). Business experience and start-up size: Buying more lottery tickets next time around? *Small Business Economics*, 43(3), 529–547.
- Coad, A., Frankish, J. S., Roberts, R. G., & Storey D. J. (2016 forthcoming). Why should banks provide entrepreneurship training seminars? *International Small Business Journal*. doi: 10.1177/0266242615593138.
- Coad, A., Frankish, J. S., Roberts, R. G., & Storey, D. J. (2013). Growth paths and survival chances: An application of Gambler's Ruin theory. *Journal of Business Venturing*, 28(5), 615–632.
- Coad, A., Frankish, J. S., Roberts, R. G., & Storey, D. J. (2015). Are firm growth paths random? A reply to "Firm growth and the illusion of randomness". *Journal of Business Venturing Insights*, 3, 5–8.
- Coad, A., & Tamvada, J. P. (2012). Firm growth and barriers to growth among small firms in India. *Small Business Economics*, 39, 383–400.
- Cox, D. R., & Snell, E. J. (1989). *The analysis of binary data* (2nd ed.). London: Chapman and Hall.
- Cumming, D. J., Ari Pandes, J., & Robinson, M. J. (2015). The role of agents in private entrepreneurial finance. *Entrepreneurship Theory and Practice*, 39(2), 345–374.
- Daunfeldt, S.-O., & Halvarsson, D. (2015). Are high-growth firms one-hit wonders? Evidence from Sweden. *Small Business Economics*, 44, 361–383.

- Davila, A., Foster, G., He, X., & Shimizu, C. (2015). The rise and fall of start-ups: Creation and destruction of revenue and jobs by young companies. *Australian Journal of Management*, 40(1), 6–35. doi:10.1177/0312896214525793.
- Dencker, J. C., Gruber, M., & Shah, S. K. (2009). Pre-entry knowledge, learning and the survival of new firms. *Organization Science*, 20(3), 516–537.
- Denrell, J., Fang, C., & Liu, C. (2015). Chance explanations in the management sciences. *Organization Science*, 26(3), 923–940.
- Frankish, J. S., Roberts, R. G., Coad, A., Spears, T. C., & Storey, D. J. (2012). Do entrepreneurs really learn? Or do they just tell us that they do? *Industrial and Corporate Change*, 22(1), 73–106.
- Frankish, J. S., Roberts, R. G., & Storey, D. J. (2010). Enterprise profiles in deprived areas: are they distinctive? *International Journal of Entrepreneurship and Small Business*, 9(2), 127–142.
- Frankish, J. S., Roberts, R. G., & Storey, D. J. (2011). Enterprise: A route out of disadvantage and deprivation? In A. Southern (Ed.), *Enterprise, deprivation and social exclusion: The role of small business in addressing social and economic inequalities*. London: Routledge.
- Fraser, S. (2005). *Finance for small and medium-sized enterprises*. CSME, University of Warwick, UK.
- Gibrat, R. (1931). Les inégalités économiques: applications, aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc.: d'une loi nouvelle la loi de l'effet proportionnel. Recueil Sirey, Paris.
- Gilbert, B. A., McDougall, P. P., & Audretsch, D. B. (2006). New venture growth: A review and extension. *Journal of Management*, 32(6), 926–950.
- Gimeno, J., Folta, T. B., Cooper, A. C., & Woo, C. Y. (1997). Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly*, 42(4), 750–783.
- Greene, F. J. (2009). Assessing the impact of policy interventions: The influence of evaluation methodology. *Environment and Planning C*, 27(2), 216–229.
- Hamilton, B. H. (2000). Does entrepreneurship pay? An empirical analysis of self-employment. *Journal of Political Economy*, 108(3), 604–631.
- Harada, N. (2007). Which firms exit and why? An analysis of small firm exits in Japan. *Small Business Economics*, 29, 401–414.
- Headd, B. (2003). Redefining business success: Distinguishing between closure and failure. *Small Business Economics*, 21, 51–61.
- Henderson, A. D., Raynor, M. E., & Ahmed, M. (2012). How long must a firm be great to rule out chance? Benchmarking sustained superior performance without being fooled by randomness. *Strategic Management Journal*, 33, 387–406.
- Hmieleski, K. M., & Baron, R. A. (2009). Entrepreneurs' optimism and new venture performance: A social cognitive perspective. *Academy of Management*, 52(3), 473–488.
- Ijiri, Y., & Simon, H. (1964). Business firm growth and size. *American Economic Review*, 54(2), 77–89.
- Jenkins, S. P. (1995). Practitioners corner: Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57(1), 129–138.
- Klotz, A. C., Hmieleski, K., Bradley, B. H., & Busenitz, L. W. (2014). New venture teams: A review of the literature and roadmap for future research. *Journal of Management*, 40(1), 226–255.
- Le Mens, G., Hannan, M. T., & Pólos, L. (2011). Founding conditions, learning, and organizational life chances: Age dependence revisited. *Administrative Science Quarterly*, 56(1), 95–126.
- Levie, J., & Lichtenstein, B. B. (2010). A terminal assessment of stages theory. Introducing a dynamic states approach to entrepreneurship. *Entrepreneurship Theory and Practice*, 34(2), 317–350.
- Levinthal, D. (1991). Random walks and organizational mortality. *Administrative Science Quarterly*, 36(3), 397–420.
- Lotti, F., Santarelli, E., & Vivarelli, M. (2009). Defending Gibrat's Law as a long-run regularity. *Small Business Economics*, 32, 31–44.
- Miller, C. C., Washburn, N. T., & Glick, W. H. (2013). The myth of firm performance. *Organization Science*, 24(3), 948–964.
- Mole, K. F., Hart, M., Roper, S., & Saal, D. S. (2011). Broader or deeper? Exploring the most effective intervention profile for public small business support. *Environment and Planning A*, 43(1), 87–105.
- Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), 691–692.
- Ongena, S., & Smith, D. C. (2000). What determines the number of bank relationships? Cross-country evidence. *Journal of Financial Intermediation*, 9(1), 26–56.
- Parker, S. C. (2009). *The economics of entrepreneurship*. Oxford: Oxford University Press.
- Persson, H. (2004). The survival and growth of new establishments in Sweden, 1987–1995. *Small Business Economics*, 23(5), 423–440.
- Phelps, R., Adams, R., & Bessant, J. (2007). Life cycles of growing organizations: A review with implications for knowledge and learning. *International Journal of Management Reviews*, 9(1), 1–30.
- Pons Rotger, G., Gørtz, M., & Storey, D. J. (2012). Assessing the effectiveness of guided preparation for new venture creation and performance: Theory and Practice. *Journal of Business Venturing*, 27(4), 506–521.
- Ryder, N. B. (1965). The cohort as a concept in the study of social change. *American Sociological Review*, 30(6), 843–861.
- Santarelli, E., & Vivarelli, M. (2007). Entrepreneurship and the process of firms' entry, survival and growth. *Industrial and Corporate Change*, 16(3), 455–488.
- Solomon, G. T., Bryant, A., May, K., & Perry, V. (2013). Survival of the fittest: Technical assistance, survival and growth of small businesses and implications for public policy. *Technovation*, 33, 292–301.
- Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., et al. (1996). Scaling behavior in the growth of companies. *Nature*, 379, 804–806.
- Storey, D. J. (1994). The role of legal status in influencing bank financing and new firm growth. *Applied Economics*, 26(2), 129–136.
- Storey, D. J. (2011). Optimism and chance: The elephants in the entrepreneurship room. *International Small Business Journal*, 29(4), 303–321.

- Storey, D. J., Keasey, K., Watson, R., & Wynarczyk, P. (1987). *The performance of small firms: Profits, jobs and failures*. London: Croom Helm.
- Syversen, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2), 326–365.
- Tornqvist, L., Vartia, P., & Vartia, Y. O. (1985). How should relative changes be measured? *American Statistician*, 39(1), 43–46.
- Voordeckers, W., & Steijvers, T. (2006). Business collateral and personal commitments in SME lending. *Journal of Banking and Finance*, 30(11), 3067–3086.
- Wennberg, K., Wiklund, J., Detienne, D. R., & Cardon, M. S. (2010). Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing*, 25, 361–375.
- Wiklund, J., Baker, T., & Shepherd, D. (2010). The age-effect of financial indicators as buffers against the liability of newness. *Journal of Business Venturing*, 25(4), 423–437.
- Yang, T., & Aldrich, H. E. (2012). Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing*, 27(4), 477–492.
- Zimmerman, M. A., & Zeltz, G. Z. (2002). Beyond survival: Achieving new venture growth by achieving legitimacy. *Academy of Management Review*, 27(3), 414–431.